

Energy Efficient Decentralized Detection

Yasaman Keshtkarjahromi, Rashid Ansari, and Ashfaq Khokhar

School of Electrical and Computer Engineering

University of Illinois at Chicago (UIC)

Chicago, Illinois 60607

Email: (ykesht2, ransari, ashfaq)@uic.edu

Abstract—In Wireless Sensor Networks (WSNs), event detection problem has been formulated as a binary hypothesis testing problem, and past work is largely focused on problem formulations that assume sensor nodes in a parallel configuration, where individual hard or soft decision is computed at each sensor node and transmitted directly to a fusion node. In such a configuration, sensor nodes farther away from the fusion node use more power to transmit their decisions. In this paper, we investigate a hierarchical configuration of the sensor nodes. In our proposed scheme, each sensor's decision is made aiming at minimizing the probability of error in the fusion node while imposing constraints on the energy consumption for information transmission. The solution is based on optimally choosing the bit allocation among the sensors and the thresholds of the decision rules. Simulation results show significant improvement in the case of the proposed hierarchical configuration compared with parallel configuration for different size networks, especially in larger networks. For example, in a 33-node network, for a fixed total number of transmitted bits, 71% less energy is consumed, and for a fixed energy budget, 146% more bits are transmitted, and 46% higher computed Chernoff information is available at the fusion node.

I. INTRODUCTION

Decentralized detection, cast as a hypothesis testing problem, involves making noisy observations at sensor nodes, locally quantizing these observations based on some decision rules, and then sending the quantized data to the fusion node for the final decision. For binary hypothesis, the goal is to decide between two states, H_0 and H_1 in the fusion node with minimum probability of error. Although the final decision is binary, the decisions at each sensor do not need to be binary. In other words, the quantization levels at the sensors may be more than two. A large body of research exists on *parallel* decentralized detection configuration, where each sensor node sends its quantized information directly to the fusion node. Decentralized detection was first introduced by Tenney and Sandell in [1]. It was extended by Tsitsiklis [2], Varshney [3], Viswanathan and Varshney [4] and Blum et al. [5]. Chamberland and Veeravali [6] and [7] investigate the problem of decentralized detection in sensor network applications, by considering resource constraints, such as spectral bandwidth, processing power, and cost, assuming that the information from the sensor nodes to the fusion node is transmitted over a wireless channel.

The major drawback of the parallel configuration is the large amount of energy needed for transmitting the quantized sensors' data. In such a configuration, nodes farther away from

the fusion node spend more energy to send their information directly to the fusion node. On the other hand, sending the quantized data of each sensor node through multiple hops significantly reduces energy consumption at the cost of increased total delay. Ferrari et. al. [8] have investigated the effect of uniform and non-uniform clustering on the probability of decision error at the fusion node, assuming each sensor node makes a binary decision about the state, and a majority-like rule is used at each intermediate fusion node.

In this paper, we study the problem of decentralized event detection formulated as a binary hypothesis testing. We introduce a hierarchical configuration of the sensor nodes based on the notion of localized clustering. Note that while hierarchical clustering concept has been applied to data gathering protocols in WSNs [9], it has not been used to formulate hypothesis testing for event detection in WSNs. In our current work, we assume that the observations at the sensors are identically and independently distributed. The number of intermediate hops for each sensor node to send its information to the fusion node is optimized as a tradeoff between the energy spent on the transmitter amplifiers (depends on the transmission distances) and the energy spent on radio electronics (does not depend on transmission distances). The decision rules and the number of quantization levels at each sensor are determined to maximize the amount of information in the fusion node, under the constraint of total energy consumed. Our simulation results show that, on average, we were able to allocate twice as many bits (keeping the total consumed energy fixed) or spend half as much energy (allocating the same number of bits to all sensor nodes) in the case of hierarchical configuration compared with parallel configuration; more improvements were shown for larger network sizes. This performance improvement is achieved at the cost of increase in the total transmission delay (which is on the order of milliseconds, thus suitable for most of the WSN applications).

II. DECENTRALIZED DETECTION PROBLEM

Let us assume $\mathbf{Y} = [Y_1, Y_2, \dots, Y_L]$ denotes a sequence vector of measurements observed over all sensor nodes, such that $Y_j = [y_j^1, y_j^2, \dots, y_j^T]$ represents the measurements at sensor node j , $1 \leq j \leq L$, and y_j^t denotes a single instance of measurement at sensor node j at time instance t , $1 \leq t \leq T$. Let us further assume that the sequence of observations at each sensor node j , has the probability density function of $f_{Y|H}(\mathbf{Y}|H_i)$, $i = 0, 1$. Each sensor quantizes its observation

according to the decision rule $\gamma_l^t : y_l^t \rightarrow u_l^t$, and then sends the quantized information to the fusion node for final decision about the state according to the decision rule $\gamma_0 : \mathbf{U} = [U_1, U_2, \dots, U_L] \rightarrow u_0$.

The goal in decentralized detection is to estimate the state in the fusion node with minimum probability of error. $\alpha = p(u_0 = H_1|H_0)$ is the probability of error, when the actual state in the environment is H_0 , while the decision of the fusion node is H_1 . Similarly, $\beta = p(u_0 = H_0|H_1)$ is the probability of error, when the actual state is H_1 , while the decision of the fusion node is H_0 . For hypothesis testing two different formulations have been used, Bayesian and Neyman-Pearson formulations.

Bayesian Formulation: In Bayesian formulation of binary hypothesis testing, a probability is assigned to H_0 and H_1 and the goal is to minimize the probability of error, $p_e = \pi_0\alpha + \pi_1\beta$, in which π_0 is the a priori probability of state H_0 and $\pi_1 = 1 - \pi_0$ is the probability of state H_1 . The achievable upper bound for the error is given by ([7]):

$$\lim_{T \rightarrow \infty} 1/T \log p_e^{(T)} \leq \log \left(\sum_{\mathbf{u}} p(\mathbf{u}|H_0)^s p(\mathbf{u}|H_1)^{1-s} \right), \quad (1)$$

\mathbf{u} is the sequence of all decision rules available at the fusion node and the inequality is true for all values of $0 \leq s \leq 1$. To minimize the probability of error, we search for the decision rules that minimize the upper bound or equivalently maximize the Chernoff information at the fusion node.

Neyman-Pearson Formulation: In Neyman-Pearson Formulation, a constraint is imposed on one of the error probabilities, α , and the goal is to minimize the other probability of error, β :

$$\text{minimize } \beta(\epsilon), \text{ subject to } 0 < \alpha < \epsilon < 1/2 \quad (2)$$

$\gamma_1, \gamma_2, \dots, \gamma_L$

According to Stein's lemma ([10]), we have:

$$\lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} 1/T \log \beta(\epsilon)^{(T)} = -D(p(\mathbf{u}|H_0)||p(\mathbf{u}|H_1)), \quad (3)$$

in which $D(a||b)$ is the relative entropy of a with respect to b or Kullback-Leibler divergence. Therefore, for minimizing the error, we should maximize $D(p(\mathbf{u}|H_0)||p(\mathbf{u}|H_1))$, which is equivalent to:

$$D(p(\mathbf{u}|H_0)||p(\mathbf{u}|H_1)) = \sum_{\mathbf{u}} p(\mathbf{u}|H_0) \log \left(\frac{p(\mathbf{u}|H_0)}{p(\mathbf{u}|H_1)} \right), \quad (4)$$

III. PROPOSED HIERARCHICAL DECENTRALIZED DETECTION CONFIGURATION

Assuming a fixed total energy budget, E , over all the sensor nodes, the goal is to minimize the probability of detection error at the fusion node. This can be achieved by maximizing the amount of information transmitted to the fusion node. Most of the existing research related to decentralized detection in WSN has assumed parallel configuration of nodes for communicating local decisions to the fusion node [6], [7]. In such a configuration, depicted in Fig. 1, the quantized information from the sensor nodes is sent directly to the fusion node.

Assuming the same radio model as in [11] to represent the communication between any pair of nodes, the consumed total energy for transmitting ($E_{(Tx)}$) and receiving ($E_{(Rx)}$) M bits from one node to another, over a distance of r , is equal to:

$$E_{(Tx)} = \varepsilon_{amp} r^2 M + E_{elec} M, E_{(Rx)} = E_{elec} M,$$

where E_{elec} is the energy dissipated to run the transmitter or receiver circuitry and ε_{amp} is the energy needed by the transmitter amplifier for transmitting one bit of information over a unit distance. Therefore nodes farther away from the fusion node will spend more energy to send the same number of bits compared with the nodes that are closer to the fusion node. We assert that this approach is inefficient in terms of energy and information quality.

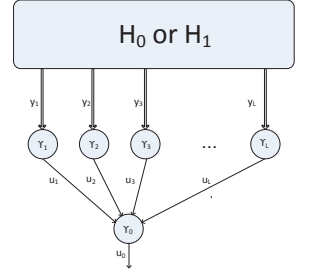


Fig. 1. Parallel Decentralized Detection Configuration

A. Hierarchical Decentralized Detection Configuration

In a hierarchical configuration, depicted in Fig. 2, sensors farther from the fusion node send their quantized information through multiple sensors closer to the fusion node, which are referred to as the intermediate fusion nodes. As an example, at the i th cluster level, which is composed of L_i sensor nodes, the nodes' quantized data is sent simultaneously to their related intermediate fusion nodes; the output of each intermediate fusion node is either the concatenation of its inputs and its quantized observation or a compressed version of them. For instance, the output of sensor node $L_1 + 1$ is equal to $u_{1,2,\dots,L_1,L_1+1} = [u_1, u_2, \dots, u_{L_1}, u_{L_1+1}]$ without compression and $u_{1,2,\dots,L_1,L_1+1} = [u_{1c}, u_{2c}, \dots, u_{L_1c}, u_{(L_1+1)c}]$ with compression, where u_{ic} represents compressed u_i . In this paper we assume no compression in the intermediate fusion nodes. Sending information via multiple hops allows for dividing the distance into shorter intervals and then summing up the energies required for transmitting and receiving in each hop. In this way, energy required for transmitter amplifiers, which is proportional to r^2 , is less compared with sending information directly to the final destination.

In our proposed hierarchical configuration, the number of intermediate fusion nodes is optimized to minimize the total energy; if we choose too many intermediate fusion nodes between a sensor node and the fusion node, then their intermediate distances from each other decrease, which results in less energy spent on the transmitter amplifiers but more energy spent on the transmitter and receiver circuitries. On the other hand, by choosing too few intermediate fusion nodes, the energy spent on the transmitter and receiver circuitries is less and the energy spent on the transmitter amplifiers is more due to larger distances between adjacent intermediate fusion nodes.

For the delay analysis of our hierarchical configuration, we assume a simple scenario in which the message from each sensor node fits a packet. The delay for transmitting each packet

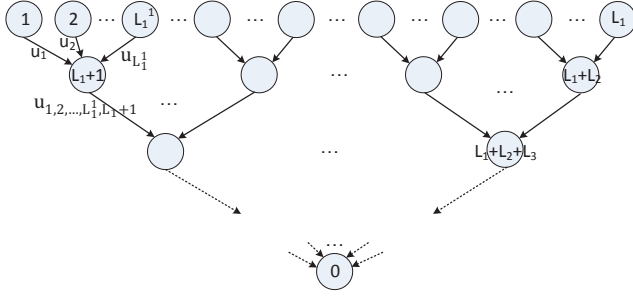


Fig. 2. Generalized Hierarchical Decentralized Detection Configuration

consists of three parts; s : sending delay for preparing a packet and putting it out on the communication link interface, c : communication delay, and r : receiving delay for processing a received packet from the communication link interface ([12]). In the proposed hierarchical configuration, delay depends on the number of cluster levels. The total delay is the sum of the delays at each cluster level, which equals to $s + c + r * L_k^*$ for the cluster level k with L_k^* as the maximum number of sensors' data received sequentially. Therefore, the total delay for K cluster levels is given by:

$$Delay = \sum_{k=1}^K (s + c + r * L_k^*) = K(s + c) + r \sum_{k=1}^K L_k^*. \quad (5)$$

Note that in the parallel configuration, the total delay for transmitting the data from L sensors to the fusion node is $s + c + r * L$. This is due to the fact that while the sensors can send their data simultaneously, the fusion node can only receive the sensors' data sequentially.

B. Determining Decision Rules

The accuracy of estimation at the fusion node depends on the amount of information the sensor nodes relay to the fusion node, which increases with an increase in the number of bits allocated to each sensor node to quantize its observation. So, both the probability of error at the fusion node and the energy consumed for transmitting the sensors' decisions to the fusion node are functions of the number of quantization levels at each sensor node. Due to the asymmetry of the location of the nodes in hierarchical configuration, allocating the same number of bits to all of the sensor nodes is not energy efficient. In this section, we solve the problem of bit allocation to the sensor nodes in order to minimize both the probability of error and the total energy consumed in transmitting the sensors' information to the fusion node.

1) *Bayesian Formulation*: Here, we assume that the observations of the sensors are conditionally independently distributed given each state: $p(\mathbf{u}|H_i) = \prod_{l=1}^L p(u_l|H_i)$ and sensor node l quantizes its observations with the number of quantization levels equal to Q_l . The Chernoff information of

the fusion node is given by ([10]):

$$\begin{aligned} C_0 &= - \min_{0 \leq s \leq 1} [\log(\sum_{\mathbf{u}} p(\mathbf{u}|H_0)^s p(\mathbf{u}|H_1)^{1-s})] \\ &= - \min_{0 \leq s \leq 1} (\log(\prod_{l=1}^L (\sum_{u_l=1}^{Q_l} p(u_l|H_0)^s p(u_l|H_1)^{1-s}))) \quad (6) \\ &\leq \sum_{l=1}^L (- \min_{0 \leq s \leq 1} (\log(\sum_{u_l=1}^{Q_l} p(u_l|H_0)^s p(u_l|H_1)^{1-s}))). \end{aligned}$$

The contribution of each sensor's decision to C_0 is equal to:

$$C_l = - \min_{0 \leq s \leq 1} (\log(\sum_{u_l=1}^{Q_l} p(u_l|H_0)^s p(u_l|H_1)^{1-s})), \quad (7)$$

C_l increases with an increase in the number of quantization levels, Q_l . Therefore, an upper bound for it is:

$$C_l < C_l^* = \lim_{Q_l \rightarrow \infty} C_l \quad (8a)$$

$$= - \min_{0 \leq s \leq 1} (\log(\int f(y_l|H_0)^s f(y_l|H_1)^{1-s} dy_l)). \quad (8b)$$

From the above equation, if the observations at all sensor nodes are identically distributed, C_l^* will be the same for all of the sensor nodes ($l = 1, 2, \dots, L$). However, allocating equal number of bits to the sensors is not energy efficient. In the hierarchical configuration, depending on the hierarchy structure, sensor nodes' quantized data may traverse different number of hops to reach the fusion node, thereby consuming different amounts of energy for transmitting their decisions. Therefore, it is obvious that under the constraint of total consumed energy, more bits may be allocated to the nodes that consume less energy for transmitting their information to the fusion node. In other words, the nodes farther from the fusion node should be allotted less number of bits and nodes closer should be assigned more bits. In the parallel configuration, we expect less total number of bits assigned to each node as more energy will be consumed on the transmission amplifiers. Assuming the same total energy available, in the hierarchical configuration, we expect more total number of bits assigned to the sensor nodes. This results in more detailed information reaching the fusion node, thus less probability of error in the hypothesis testing.

Assume, l th route consisting of node l (the first node in the route), $N_l - 1$ intermediate fusion nodes, and the fusion node (node 0), is used for transmitting node l 's quantized data (Fig. 3). d_{ij} is the distance between two adjacent nodes i and j in the route. The energy consumed in transmitting sensor l 's quantized data via route

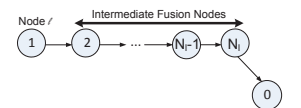


Fig. 3. l th Route Consisting of Node l (the first node in the route), $N_l - 1$ Intermediate Fusion Nodes and the Fusion Node

l is:

$$E_l = \sum_{n=1}^{N_l} E_{(Tx)n} + E_{(Rx)n} \quad (9a)$$

$$= M_l(2N_l E_{elec} + \varepsilon_{amp} d_{N_l 0}^2 + \sum_{n=1}^{N_l-1} \varepsilon_{amp} d_{n(n+1)}^2), \quad (9b)$$

where, $E_{(Tx)n}$ and $E_{(Rx)n}$, are the energies dissipated for transmitting and receiving from node n to the next node in the rout and M_l is the number of bits allocated to node l , $M_l = \log_2(Q_l)$.

For bit allocation among the sensors, our minimization problem is aimed at determining the decision rules of the sensors that maximize the Chernoff information subject to a constraint of total consumed energy. The decision rule at each sensor node is completely identified by its number of quantization levels and decision regions. According to $C_0 \leq \sum_{l=1}^L C_l$, the minimization problem is formulated as:

$$\text{maximize}_{\gamma_1, \gamma_2, \dots, \gamma_L} \sum_{l=1}^L C_l, \text{ subject to } \sum_{l=1}^L E_l \leq E \quad (10)$$

To solve this optimization problem, first we find decision regions for each sensor, which maximizes its Chernoff information, C_l , for different values of quantization levels, Q_l . The decision rule used at each sensor node for quantization, is Maximum Likelihood Ratio (MLR) test; for a fixed number of quantization levels, Q_l , the optimum decision regions are obtained accordingly: $\gamma_l^{\text{opt}}(Q_l) = \arg \max_{\gamma_l(Q_l)} C_l(Q_l)$, where,

$$C_l(Q_l) = - \min_{0 \leq s \leq 1} \left(\log \left(\sum_{u_l=1}^{Q_l} \left(\int_{y_l \in \gamma_l^{-1}(u_l)} f(y_l|H_0) dy_l \right)^s \left(\int_{y_l \in \gamma_l^{-1}(u_l)} f(y_l|H_0) dy_l \right)^{1-s} \right) \right), \quad (11)$$

In the next step, we solve the optimization problem to determine $\mathbf{Q} = [Q_1, Q_2, \dots, Q_L, \dots, Q_L]$:

$$\text{maximize}_{Q_1, Q_2, \dots, Q_L} \sum_{l=1}^L C_l(Q_l), \text{ subject to } \sum_{l=1}^L E_l \leq E \quad (12)$$

2) *Neyman-Pearson Formulation*: With the assumption of conditionally independently distributed observations, we have:

$$D(p(\mathbf{u}|H_0)||p(\mathbf{u}|H_1)) = \sum_{l=1}^L \left(\sum_{u_l=1}^{Q_l} (p(u_l|H_0) \log(\frac{p(u_l|H_0)}{p(u_l|H_1)})) \right), \quad (13)$$

from which the contribution of each sensor and its upper bound are:

$$\sum_{u_l=1}^{Q_l} (p(u_l|H_0) \log(\frac{p(u_l|H_0)}{p(u_l|H_1)})) \leq \int (f(y_l|H_0) \log(\frac{f(y_l|H_0)}{f(y_l|H_1)})) dy_l. \quad (14)$$

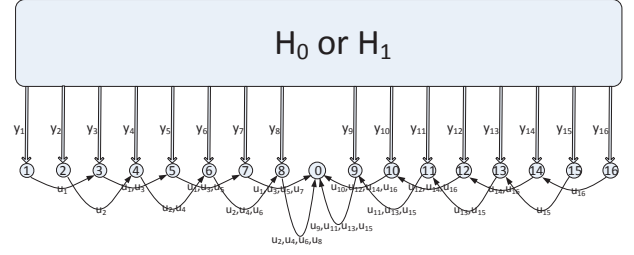


Fig. 4. Optimum Hierarchical Configuration for Forwarding Sensor's Quantized Data in a 17 Sensor Node Network

Using the procedure similar to what is discussed in the previous section for the Bayesian formulation, the optimum decision regions are obtained from:

$$\gamma_l^{\text{opt}}(Q_l) = \arg \max_{\gamma_l(Q_l)} \sum_{u_l=1}^{Q_l} (p(u_l|H_0) \log(\frac{p(u_l|H_0)}{p(u_l|H_1)})), \quad (15)$$

and the optimization problem is:

$$\begin{aligned} & \text{maximize}_{Q_1, Q_2, \dots, Q_L} \sum_{l=1}^L \sum_{u_l=1}^{Q_l} (p(u_l|H_0) \log(\frac{p(u_l|H_0)}{p(u_l|H_1)})) \\ & \text{subject to } \sum_{l=1}^L E_l \leq E \end{aligned}$$

IV. SIMULATION RESULTS

The performance of the proposed hierarchical configuration of nodes for the decentralized detection problem is evaluated by simulating in MATLAB and compared with the parallel configuration in terms of energy cost, information quality and delay. We simulated Gaussian random variables of observations in the sensors; $f(y|H_0) = \mathcal{N}(-1, 1)$ and $f(y|H_1) = \mathcal{N}(1, 1)$ are considered for determining optimized decision rules within the sensor nodes. For simplicity, we assume a single dimensional WSN field, where the nodes are equispaced on a straight line (see Fig. 4 for a 17 sensor node network), however the proposed method can be applied to any arrangement of sensors in the network. Let us assume that the distance between each adjacent node is d by defining the sensor node in the middle, i.e. node 0, as the fusion node. The values for ε_{amp} and E_{elec} used in the simulations are: $\varepsilon_{amp} = 100pJ/bit/m^2$, $E_{elec} = 50nJ/bit$.

A. Hierarchical vs Parallel Configuration

We have used different values for inter-node distances, d , in our simulations and the results show very similar trends. In this paper we report only the results for $d = 20m$. Fig. 5 shows the total energy for transmitting one bit of quantized information at each node in both the parallel and hierarchical configurations. The data in this figure is applicable to different size networks, from 3 up to 33 nodes. The starting point of each curve in Fig. 5 (i.e number of intermediate fusion nodes = 0), corresponds to the parallel configuration and the other points (number of intermediate fusion nodes > 0) are related

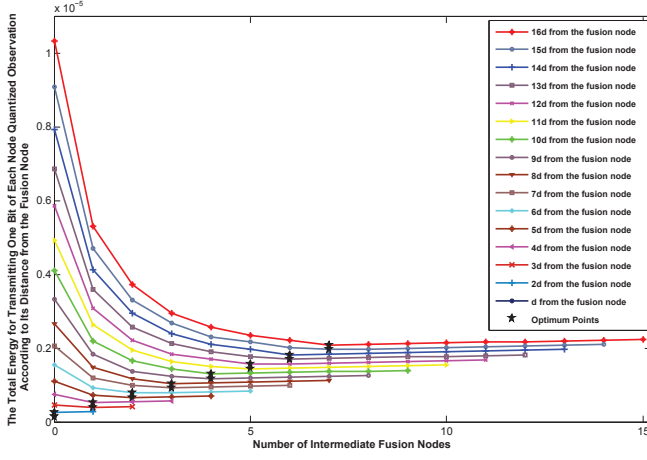


Fig. 5. The Total Energy Consumed for Transmitting One Bit of Information by Nodes at Different Distances from the Fusion Node

to different number of cluster levels in a hierarchical configuration, therefore simulating different hierarchical configurations. As can be seen from the figure, for each node's quantized data, the parallel configuration consumes more energy compared with the hierarchical configuration.

In the Hierarchical configuration, for each sensor node, depending on its distance from the fusion node, the optimum number of intermediate fusion nodes needs to be selected. As discussed in Section III-A, this number should be determined such that it minimizes the total required energy. The optimum points are shown by a star in Fig. 5.

In addition, the energy consumed for different nodes' quantized data is an increasing function of the node's distance from the fusion node in both parallel and hierarchical configuration; by allocating the same number of bits to all of the sensor nodes' observation, more energy is consumed to transmit the quantized data of farther nodes from the fusion node.

The total delay for the optimum hierarchical configuration in Fig. 4 (using the results of Fig. 5) consists of four parts related to the four cluster levels:

- 1) $s + c + r$ for simultaneous transmission of node 1 to 3, 2 to 4, 16 to 14 and 15 to 13
- 2) $s + c + r$ for simultaneous transmission of node 3 to 5, 4 to 6, 14 to 12 and 13 to 11
- 3) $s + c + r$ for simultaneous transmission of node 5 to 7, 6 to 8, 12 to 10 and 11 to 9
- 4) $s + c + 4 * r$ for simultaneous transmission of node 7, 8, 9 and 10 to the fusion node.

In parallel configuration, the delay is equal to $s + c + 16 * r$. Therefore, if we assume sending and receiving overheads to be negligible relative to the cost of actual communication, the total delay in the parallel configuration, which is $s + c + 16 * r \simeq c$, is approximately one forth of the delay in the hierarchical configuration, $4 * s + 4 * c + 7 * r \simeq 4 * c$.

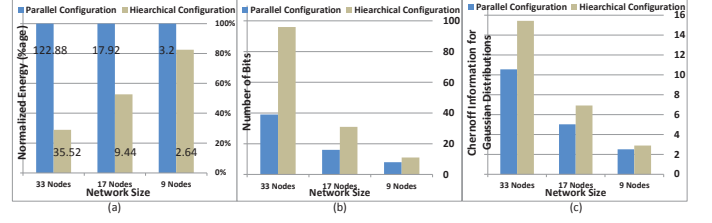


Fig. 6. Comparison of the (a) Total Energy Consumed, (b) Total Number of Bits Received at the Fusion Node, and (c) Chernoff Information in the Fusion Node, for 9, 17 and 33 Sensor Node Network

TABLE I
OPTIMIZED DECISION RULES AT EACH SENSOR NODE FOR DIFFERENT VALUES OF Q , GAUSSIAN DISTRIBUTION OF BAYESIAN FORMULATION

Q	thresholds	$C_{opt}(Q)$
2	[0]	0.3137
4	[-1 0 1]	0.4399
8	[-1.8 -1.1 -0.5 0 0.5 1.1 1.8]	0.4824

B. Effect of Scaling the Network Size

Fig. 6(a) shows total energy consumed by hierarchical versus parallel configurations when one bit is allocated to the quantized data of each node. We see that larger networks show significantly more improvement in terms of energy savings. The results shown in Figs. 6(b) and 6(c) were obtained under a fixed total energy constraint for each of the sensor networks and by applying the method discussed in section V. Fig. 6(b) shows that the total number of bits received by the fusion node was improved significantly for the case of hierarchical configuration; again higher percentage improvements were shown for larger networks. Fig. 6(c) shows that the same bit allocation, results in significant improvements in terms of the fusion node's Chernoff information (using the results from section IV-C).

C. Effect of Quantization

First, we need to determine the decision rules of all the nodes for different values of Q , the number of quantization levels. From the fact that the Gaussian distributions satisfy the monotone likelihood ratio property, we determine the thresholds of quantization at each sensor node accordingly, for different values of Q . Since the probability distributions at all the nodes are identical, the minimization problem for all of the nodes is identical and is expressed as:

$$C_{opt}(Q) = \max_{t_1, t_2, \dots, t_{Q-1}} C_l(Q), \quad (16)$$

and

$$\max_{t_1, t_2, \dots, t_{Q-1}} \sum_{u_l=1}^{Q_l} (p(u_l|H_0) \log(\frac{p(u_l|H_0)}{p(u_l|H_1)})) \quad (17)$$

for Bayesian and Neyman-Pearson formulations respectively. Here $t_i, i = 1, 2, \dots, Q - 1$ are the thresholds that have the property $t_1 < t_2 < \dots < t_{Q-1}$ under the monotone likelihood ratio. The optimum thresholds for different values

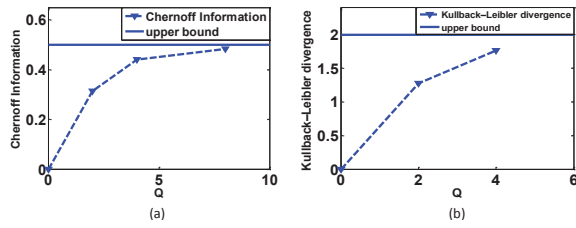


Fig. 7. Optimized (a) Chernoff Information, and (b) Kullback-Leibler Divergence, for Different Values of Quantization Levels

of quantization levels were obtained by simulation and are listed in Tables I and II with four-digit accuracy of Chernoff information and Kullback-Leibler divergence. As the simulation results show, Fig. 7, the Chernoff information and Kullback-Leibler divergence increase with the increase in the number of quantization levels and get closer to the upper bounds 0.5 and 2.0 respectively.

V. DISCUSSION

Under the constraint of fixed total energy, solving the optimum bit allocation problem using known optimization methods such as Lagrangian method does not result in an explicit solution. Therefore, an algorithm-oriented method for Gaussian distributions is proposed, which is an iterative approach for increasing the number of bits allocated to the sensors. This algorithm is based on the fact that, allocating minimum number of bits to each node (two bits), results in considerable amounts of Chernoff Information and Kullback-Leibler divergence, more than 60% of the upper bound. In addition, as shown in Fig. 7, the amounts of Chernoff information and Kullback-Leibler divergence increase rapidly at first with an increase in the number of allocated bits and then increase slowly:

- 1) Change the numbering of the sensor nodes in the order of increasing energy required for transmitting one bit of information, i.e. $E_1 \leq E_2 \leq \dots \leq E_L$
- 2) Set $M_i = \log_2(Q_i) = 0, i = 1, 2, \dots, L$ with the computed consumed energy equal to zero.
- 3) While the computed consumed energy is less than the constraint energy and $L \neq 0$:
 - a) $i = 1$.
 - b) While $i \leq L$:
 - i) $M_i = M_i + 1$;
 - ii) Compute the energy consumed with this bit allocation.
 - iii) If the computed consumed energy is more than the energy constraint,
 - A) $M_i = M_i - 1$.
 - B) $L = i - 1$.
 - C) $i = 0$.
 - D) Compute the energy consumed with this bit allocation.
 - iv) $i = i + 1$.

TABLE II
OPTIMIZED DECISION RULES AT EACH SENSOR NODE FOR $Q = 2, 4$, GAUSSIAN DISTRIBUTION OF NEYMAN-PEARSON FORMULATION

Q	thresholds	$\sum_{u_l=1}^{Q_l} (p(u_l H_0) \log(\frac{p(u_l H_0)}{p(u_l H_1)}))$
2	[-0.6]	1.2788
4	[-1.7 -0.7 0.3]	1.7653

- 4) Renumber the sensor nodes as they were before applying the algorithm.
- 5) Report the number of bits allocated as the optimum bit allocation among the nodes.

VI. CONCLUSION

In this paper, an energy efficient decentralized detection is studied using a hierarchical configuration of the sensor nodes. The number of intermediate hops for each sensor node is determined to achieve minimum energy consumption for transmitting the information to the fusion node. In addition, a methodology for determining the decision rules at each sensor node and the number of bits allocated are proposed to maximize the amount of information at the fusion node. We show that compared with the existing formulations based on parallel configuration, the proposed hierarchical configuration achieves better energy consumption at the cost of minor addition in the latency. We envision a hybrid system that can be dynamically adapted to the application constrains while minimizing energy usage as well as delay.

REFERENCES

- [1] R. R. Tenney and N. R. Sandell, "Detection with distributed sensors," in *Decision and Control including the Symposium on Adaptive Processes, 1980 19th IEEE Conference on*, vol. 19, Dec. 1980, pp. 433–437.
- [2] J. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing*, vol. 2, 1993, pp. 297–344.
- [3] P. Varshney, *Distributed Detection and Data Fusion*. New York: Springer-Verlag, 1996.
- [4] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors i. fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, Jan 1997.
- [5] R. Blum, S. Kassam, and H. Poor, "Distributed detection with multiple sensors i. advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, Jan 1997.
- [6] J.-F. Chamberland and V. Veeravalli, "Wireless sensors in distributed detection applications," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 16–25, May 2007.
- [7] —, "Decentralized detection in sensor networks," *Signal Processing, IEEE Transactions on*, vol. 51, no. 2, pp. 407–416, Feb 2003.
- [8] G. Ferrari, M. Martalo, and R. Pagliari, "Decentralized detection in clustered sensor networks," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 47, no. 2, pp. 959–973, April 2011.
- [9] S. Bandyopadhyay and E. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3, March-3 April 2003, pp. 1713–1723 vol.3.
- [10] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [11] S. Lindsey, C. Raghavendra, and K. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 13, no. 9, pp. 924–935, Sep 2002.

- [12] S. Hambrusch and A. Khokhar, "C3: an architecture-independent model for coarse-grained parallel machines," in *Parallel and Distributed Processing, 1994. Proceedings. Sixth IEEE Symposium on*, Oct 1994, pp. 544–551.